

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2003, 23:55:25 ; Search time 6291 Seconds

(without alignments)
9918.362 Million cell updates/sec

Title: US-09-698-781-2

Sequence: 1 tgaatgaacaatactactcat.....gaaaaaaaaaaaaaaaaa 2144

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rnd:*
- 36: em_htg_rnd:*
- 37: em_htg_vit:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2144	100.0	2144	6 AX127587	AX127587 Sequence
2	2134.8	99.6	2138	9 HSPC28	X94323 H. sapiens m
3	2133.2	98.6	2128	6 AX35634	AX35634 Sequence
4	2113.2	98.6	2128	2 HSCRSPPG	X95240 H. sapiens m
5	1485.2	69.3	151752	2 AC010779	AC010779 Homo sapi
6	1480.4	69.0	118524	9 HSDJ44216	AL121950 Homo DNA
7	515.2	24.0	1335	9 HUMPEX1A	M25532 Human testi
8	513.6	24.0	1380	9 BC022011	BC022011 Homo sapi
9	513.6	24.0	1406	9 HSCRSPP2I	X95239 H. sapiens m
10	505.8	23.6	1295	4 ECCRISP3	AJ001400 Equus cab
11	476.2	22.2	1388	10 CPUS3712	U35712 Cavia porce
12	370.8	17.3	1418	10 MURSPX1A	M25533 Mouse testl
13	367.8	17.2	1280	10 AB009662	AB009662 Rattus no
14	362.2	16.9	1432	10 AF078552	AF078552 Rattus no
15	362.2	16.9	1498	10 RATBEG	M31173 Rat epididy
16	346.4	16.2	1380	10 MUSAE61A	M92849 Mouse acidl
17	346.4	16.2	1403	10 MUSCRISPA	L05559 Mouse cyste
18	341.6	15.9	1445	10 BC011150	BC011150 Mus muscu
19	325.6	15.2	928	10 RNEBGR	X04643 Rat mRNA fo
20	325.4	15.2	495	4 ECA6632	AJ006632 Equus cab
21	284.4	13.7	453	4 BTA277708	AJ277708 Bos tauru
22	287.6	13.4	1394	10 MUSAE62A	M92850 Mouse acidl
23	287.6	13.4	1418	10 BC022573	BC022573 Mus muscu
24	287.6	13.4	1406	10 MUSCRISP8	L05560 Mouse cyste
25	254.4	11.9	256	6 AX127586	AX127586 Sequence
26	226.4	10.6	1090	5 U13619	U13619 Heloderma h
27	210.8	9.8	334	5 BTA277709	AJ277709 Bos tauru
28	205	9.6	1336	5 AF384218	AF384218 Agkistrod
29	202.4	9.4	1321	5 AF384219	AF384219 Trimeresu
30	198.6	9.3	791	5 AY093955	AY093955 Rhadodophl
31	198.6	9.3	1318	5 PMU59447	U59447 Trimeresuru
32	183	8.5	51371	9 AL353193	AL353193 Human DNA
33	182.8	8.5	169189	2 AC124908	AC124908 Equus cab
34	182.4	8.5	1309	5 AF190861	AF190861 Lepemus h
35	182	8.5	1475	4 ECA315379	AJ315379 Equus cab
36	180.8	8.4	1305	5 AF384220	AF384220 Laticauda
37	179.8	8.4	155869	2 AC023420	AC023420 Homo sapi
38	177.4	8.3	1316	5 AF159541	AF159541 Lepemus h
39	177	8.3	1281	9 HUMARP	D38451 Homo sapien
40	177	8.3	1282	9 S80310	S80310 Homo sapien
41	177	8.3	1886	9 HSCRSPIG	X95237 H. sapiens m
42	175.6	8.2	1174	4 ECA6631	AJ006631 Equus cab
43	173.4	8.1	149907	4 AC091436	AC091436 Felis cat
44	170.2	7.9	166287	9 AC026843	AC026843 Homo sapi
45	169.2	7.9	1343	9 AF123894	AF123894 Macaca mu

ALIGNMENTS

RESULT 1
AX127587
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORFANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2144)
REFERENCE
AUTHORS
TITLE

Pred. No. is the number of results predicted by chance to have a

|||||
Db 1981 TATCTTGACAAATATAGAGGGTAAATGTAAGTAACTAGAGAACTGAACTGG 2040
OY 2041 GTGAGGCTATCTAGCAATCTCTCTACTATCTTACCAATTTTCGTAAGTAAAGAA 2100
|||||
Db 2041 GTGAGGCTATCTAGCAATCTCTCTACTATCTTACCAATTTTCGTAAGTAAAGAA 2100
OY 2101 GCAATGCAAAATTAAGATATCTTGAAGAAAAA 2144
|||||
Db 2101 GCAATGCAAAATTAAGATATCTTGAAGAAAAA 2144
|||||
RESULT 2
HSPG28 2138 bp mRNA linear PRI 04-MAR-1996
LOCUS HSPG28
DEFINITION H.sapiens mRNA for SGP28 protein.
ACCESSION X94323
VERSION X94323.1 GI:1213612
KEYWORDS glycoprotein; SPG28 protein.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2138)
AUTHORS Kjeldsen, L., Cowland, J.B., Johnsen, A.H. and Borregaard, N.
TITLE SGP28, a novel matrix glycoprotein in specific granules of human
neutrophils with similarity to a human testis-specific gene product
and a rodent sperm-coating glycoprotein
JOURNAL FEBS Lett. 380 (3), 246-250 (1996)
MEDLINE 96186934
PUBMED 8601434
REFERENCE 2 (bases 1 to 2138)
AUTHORS Cowland, J.B.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1995) J.B. Cowland, Granulocyte Research
Laboratory, Dept of Hematology, National Univ. Hosp.,
Rigshospitalet L-4041, 9 Blegdamsvej, 2100 Copenhagen, DENMARK
FEATURES
source
1. 2138
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="bone marrow"
41..778
/codon_start=1
/product="SGP28 protein"
/protein_id="CAA63984.1"
/db_xref="GI:1213612"
CDS
/db_xref="SWISS-PROT:P54108"
/translation="MTLFFVLFLVAGLLSPFPAEDKDPATFALLITQYQVREIVN
KHNELRRAVSPPARNMLKMEWKEAANAKNANCNYSHPKDMTSLKGEENLYM
SSAPSSWSQAIOISWFDYNDPFGVGPKNPAVVGHTVYVYSSLYVCGNAYCPNO
KYLKYYVCOYCPAGNMANRLVPEQGPAPCASCPNCDDGLCTNCKKEDLYSNCKS
IKLTFTCKQLVDRDCKASCNCNSNY"
BASE COUNT 728 a 404 c 383 g 623 t
ORIGIN
Query Match 99.6%; Score 2134.8; DB 9; Length 2138;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

|||||
Db 181 TGAAGTGAAGAGAGAGATATCTCCCTGCGACAAAACATGCTGAAGTGAAGAACAA 240
OY 242 AGAGGCTGCGCAAAATGCCAAAAGTGGGCAACACAGTGCATTAAGACACAGTAACCC 301
|||||
Db 241 AGAGGCTGCGCAAAATGCCAAAAGTGGGCAACACAGTGCATTAAGACACAGTAACCC 300
OY 302 AAAGATGCAATGACAAAGTCTAAATGTGTGAGAAATCTCTACATGTCAAGTCCCCAG 361
|||||
Db 301 AAAGATGCAATGACAAAGTCTAAATGTGTGAGAAATCTCTACATGTCAAGTCCCCAG 360
OY 362 CTGATGCTCACAAGCAATCCAAAGCTGTTTGTATGAGTACATGATTTGACTTGTGTG 421
|||||
Db 361 CTGATGCTCACAAGCAATCCAAAGCTGTTTGTATGAGTACATGATTTGACTTGTGTG 420
OY 422 AGGGCCAAAGACTGCCAAGCAGTGGTGGACATTTACACAGTGTGTTGGTACTCTTC 481
|||||
Db 421 AGGGCCAAAGACTGCCAAGCAGTGGTGGACATTTACACAGTGTGTTGGTACTCTTC 480
OY 482 ATACCTGCTTGATGGAATGCTACTGTGCCAATCAAAAAGTCTAAATATCTACTA 541
|||||
Db 481 ATACCTGCTTGATGGAATGCTACTGTGCCAATCAAAAAGTCTAAATATCTACTA 540
OY 542 TGTTCGCCAATATTTGCTGCTGCTGTAATGGCTAATAGCTATATGTCCCTATGACA 601
|||||
Db 541 TGTTCGCCAATATTTGCTGCTGCTGTAATGGCTAATAGCTATATGTCCCTATGACA 600
OY 602 AGGAGCACTTGTGCCAGTGGCCAGATACTGACATGAGATGACATGACCAATGTTG 661
|||||
Db 601 AGGAGCACTTGTGCCAGTGGCCAGATACTGACATGAGATGACATGACCAATGTTG 660
OY 662 CAAATGAGAGATCTGTATAGTAACTGTAAAGCTTGAAGCTACATTAACCTGTAACA 721
|||||
Db 661 CAAATGAGAGATCTGTATAGTAACTGTAAAGCTTGAAGCTACATTAACCTGTAACA 720
OY 722 TCACTGCTGAGGAGCAGTGGCAAGGCAATCTGCAATGTTCAAAAGCATTTATTAAT 781
|||||
Db 721 TCACTGCTGAGGAGCAGTGGCAAGGCAATCTGCAATGTTCAAAAGCATTTATTAAT 780
OY 782 ACCGATTACACACGAGTAGGAGCTATGTAGAGAGAGTCAATATCTACTTATGATTGG 841
|||||
Db 781 ACCGATTACACACGAGTAGGAGCTATGTAGAGAGAGTCAATATCTACTTATGATTGG 840
OY 842 CATCTACTAGATTAAATATCTAGCTGAGAAATGTAGGCAATGTTGATACACATTT 901
|||||
Db 841 CATCTACTAGATTAAATATCTAGCTGAGAAATGTAGGCAATGTTGATACACATTT 900
OY 902 GATTTCAAATGTTTCTCTGATCTGCTTTTATTTTCAAAAATATTTTCATACAA 961
|||||
Db 901 GATTTCAAATGTTTCTCTGATCTGCTTTTATTTTCAAAAATATTTTCATACAA 960
OY 962 ATGGTTAAAGAAACAAATCTATAACAACAACACTTGGATTTTATATATAAATTTGT 1021
|||||
Db 961 ATGGTTAAAGAAACAAATCTATAACAACAACAACACTTGGATTTTATATATAAATTTGT 1020
OY 1022 GATTTAATTTACTGAATTAATTTAGGGTGAATTTTGAAGTGTATTTCTATATGAC 1081
|||||
Db 1021 GATTTAATTTACTGAATTAATTTAGGGTGAATTTTGAAGTGTATTTCTATATGAC 1080
OY 1082 TAAGTTCACTAAACCTGGATTTGAAGTGAATTTGTCTCAGAACAAATGTTCACAA 1141
|||||
Db 1081 TAAGTTCACTAAACCTGGATTTGAAGTGAATTTGTCTCAGAACAAATGTTCACAA 1140
OY 1142 AAAGAACATATATTTTCAATGAACCTTGGCTGTAGTTGGCTTCTAGTCCACTC 1201
|||||
Db 1141 AAAGAACATATATTTTCAATGAACCTTGGCTGTAGTTGGCTTCTAGTCCACTC 1200
OY 1202 TAAAGCTAAGCATCTTCAAAAGAGCTTTTCCCATATGCTGTCTTATTTTCACTCAT 1261
|||||
Db 1201 TAAAGCTAAGCATCTTCAAAAGAGCTTTTCCCATATGCTGTCTTATTTTCACTCAT 1260
OY 1262 CACCTCTTCCCAATCATGTGGCTGGCATCTCACAAATGAGTGAAGTGTTCCTCCT 1321
|||||

Db 1261 CACCCTTTCCCAATCATCTGGCTGGCATCCCTCACAAATGAGTTGAAGCTGTCTCCT 1320
OY 1322 AAAACAATCTGCACTTTTATTTTGGCAAAATCAATACATCTTTGAATTTTATPCTGC 1381
Db 1321 AAAACAATCTGCACTTTTATTTTGGCAAAATCAATACATCTTTGAATTTTATPCTGC 1380
OY 1382 ATAAATTTTACAGTAAGTAATGATCAAACTTCATTTTAAACCTCTCTCTCTTGACA 1441
Db 1381 ATAAATTTTACAGTAAGTAATGATCAAACTTCATTTTAAACCTCTCTCTCTTGACA 1440
OY 1442 AAATCTCTTAAAGTAAGTAATGATCAAACTTCATTTTAAACCTCTCTCTCTTGACA 1501
Db 1441 AAATCTCTTAAAGTAAGTAATGATCAAACTTCATTTTAAACCTCTCTCTCTTGACA 1500
OY 1502 GAATCTGCT 1561
Db 1501 GAATCTGCT 1560
OY 1562 GAGTATGGAAGGGAAGCACTAGTAATCTTACAGGGGAGAAAAATGACAAATGAGCTCTTC 1621
Db 1561 GAGTATGGAAGGGAAGCACTAGTAATCTTACAGGGGAGAAAAATGACAAATGAGCTCTTC 1620
OY 1622 ACCAAGTATCAAAATTAAGCTCACAGATAGTCAATTCAGATTGTGTAGATTAATC 1681
Db 1621 ACCAAGTATCAAAATTAAGCTCACAGATAGTCAATTCAGATTGTGTAGATTAATC 1680
OY 1682 TTTCTAAATTAATCATTAATCCCAATCTAATTAATGAGCTTAAACATCCAGCAATCAAGTT 1741
Db 1681 TTTCTAAATTAATCATTAATCCCAATCTAATTAATGAGCTTAAACATCCAGCAATCAAGTT 1740
OY 1742 GAAGGACATTTCTACAAATATCTCTGGGGATTTTGAAGTATCTCTCAAACTGTAAAAA 1801
Db 1741 GAAGGACATTTCTACAAATATCTCTGGGGATTTTGAAGTATCTCTCAAACTGTAAAAA 1800
OY 1802 TCATGGAATTAAGGAATCTCTGAGAAACAAATCACAGACCAATGAGACTAAGGAGACT 1861
Db 1801 TCATGGAATTAAGGAATCTCTGAGAAACAAATCACAGACCAATGAGACTAAGGAGACT 1860
OY 1862 GTGAGCCAAATGCAATGTGCTTCTTGATCAGATCCTGGAACAGAAAAAGATCAGTAATG 1921
Db 1861 GTGAGCCAAATGCAATGTGCTTCTTGATCAGATCCTGGAACAGAAAAAGATCAGTAATG 1920
OY 1922 AAAAATCGATGAATCTGGAATAGAAATCTGAGTATTTTACAGTAAGTGTGATTTCTT 1981
Db 1921 AAAAATCGATGAATCTGGAATAGAAATCTGAGTATTTTACAGTAAGTGTGATTTCTT 1980
OY 1982 AATCTTGACAATATAGAGAGGTAATGTAAATGATTAAGTAACTGAGAACTGAGG 2041
Db 1981 AATCTTGACAATATAGAGAGGTAATGTAAATGATTAAGTAACTGAGAACTGAGG 2040
OY 2042 TGAGGGCTATCTAGGAATTTCTGTACTATCTTACCAATTTTGGGTAAGTCTAAGAAAG 2101
Db 2041 TGAGGGCTATCTAGGAATTTCTGTACTATCTTACCAATTTTGGGTAAGTCTAAGAAAG 2100
OY 2102 CAATGCAAAATTAAGTAATCTTGAAGAAAAA 2139
Db 2101 CAATGCAAAATTAAGTAATCTTGAAGAAAAA 2138

RESULT 3
AX335634
LOCUS AX335634 2128 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 6143 from Patent WO0194629.
ACCESSION AX335634
VERSION AX335634.1 GI:18126353
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 Young, P. E., Augustus, M., Carter, K. C., Ehner, R., Endress, G.,
Horrikan, S., Soppet, D. R. and Weaver, Z.

TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL Patent: WO 0194629-A 6143 13-DEC-2001;
FEATURES
source Avalon Pharmaceuticals (US)
Location/Qualifiers
1. 2128
/organism="Homo sapiens"
/db_xref="Laxon:9606"
BASE COUNT 734 a 397 c 380 g 617 t
ORIGIN
Query Match 98.6%; Score 2113.2; DB 6; Length 2128;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2115; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 27 CTGGAACCACTGCAATGCAATGATTTCCAGTGGTGTCTGCTGGGCTGCTT 86
Db 1 CTGGAACCACTGCAATGCAATGATTTCCAGTGGTGTCTGCTGGGCTGCTT 60
OY 87 CCATCTTTCCAGCAATGAAGATAGGATCCCGCTTTTACTGCTTTGTAAACCCCA 146
Db 61 CCATCTTTCCAGCAATGAAGATAGGATCCCGCTTTTACTGCTTTGTAAACCCCA 120
OY 147 ACACAAGTGAAGGAGAGATTTGTGAATAGCAATGAAGTGAAGAGAGAGAGTCTCC 206
Db 121 ACACAAGTGAAGGAGAGATTTGTGAATAGCAATGAAGTGAAGAGAGAGAGTCTCC 180
OY 207 CTGCGCAGAACATCTGAGATGGAATGGAACCAAGAGGCTGCGCAATGCCCAAG 266
Db 181 CTGCGCAGAACATCTGAGATGGAATGGAACCAAGAGGCTGCGCAATGCCCAAG 240
OY 267 TGGGCAACCACTGCAATTTACAGACAGTAACCCCAAGAGATGCAATGCAATGCA 326
Db 241 TGGGCAACCACTGCAATTTACAGACAGTAACCCCAAGAGATGCAATGCAATGCA 300
OY 327 TGTGTGAGAAATCTTACATGTCAGAGTCCCGGCTGCTGCTGCTGCTGCTGCTGCT 386
Db 301 TGTGTGAGAAATCTTACATGTCAGAGTCCCGGCTGCTGCTGCTGCTGCTGCTGCT 360
OY 387 TGTGTGAGAAATCTTACATGTCAGAGTCCCGGCTGCTGCTGCTGCTGCTGCTGCT 446
Db 361 TGTGTGAGAAATCTTACATGTCAGAGTCCCGGCTGCTGCTGCTGCTGCTGCTGCT 420
OY 447 GTTGACATTAATACAGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 506
Db 421 GTTGACATTAATACAGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 480
OY 507 TACTGTCCCAATCAAAAGTTCTAAATTAATCTACTATGTTGGCAATATGCTCTGCTG 566
Db 481 TACTGTCCCAATCAAAAGTTCTAAATTAATCTACTATGTTGGCAATATGCTCTGCTG 540
OY 567 AATTGGGTAATAGACTATATGCTCTTATGAACAAGAGGACCTGTTGCCAGTTGCCA 626
Db 541 AATTGGGTAATAGACTATATGCTCTTATGAACAAGAGGACCTGTTGCCAGTTGCCA 600
OY 627 GATAACTGTGAGATGAGATATGACCAATGTTGCCAATGAGATCTATAGTAAC 686
Db 601 GATAACTGTGAGATGAGATATGACCAATGTTGCCAATGAGATCTATAGTAAC 660
OY 687 TGTAAAGTTGTAAGCTACATTAACCTGTAACATCACTGTTGCTGAGAGAGTGGCAAG 746
Db 661 TGTAAAGTTGTAAGCTACATTAACCTGTAACATCACTGTTGCTGAGAGAGTGGCAAG 720
OY 747 GCATCTGCAATTTGTTCACAACAGATTTATTAATAGCATTAACAGAGAGTGGCTA 806
Db 721 GCATCTGCAATTTGTTCACAACAGATTTATTAATAGCATTAACAGAGAGTGGCTA 780
OY 807 TGTAGAGAGAGTCAATTAATCTACTAGATTGGCATCTACTAGATTGAATTAACATACT 866
Db 781 TGTAGAGAGAGTCAATTAATCTACTAGATTGGCATCTACTAGATTGAATTAACATACT 840
OY 867 AGCTAGAGAAATTTGGAGATGTTGATACACATTTGATTTCAATGTTTCTTCTGAT 926
Db 847 AGCTAGAGAAATTTGGAGATGTTGATACACATTTGATTTCAATGTTTCTTCTGAT 900

```

Db 841 AGCTGAAATTTAGCATGTTGATACATTTGATTTCAATGTTTTCTTCTGAT 900
Qy 927 CTGCTTTTATTTACAAAAATTTTTCATACAAATGGTTAAAGAAACAAATCTAT 966
Db 901 CTGCTTTTATTTACAAAAATTTTTCATACAAATGGTTAAAGAAACAAATCTAT 960
Qy 987 AACACAACTTTGGATTTTATATATTAACCTTGTGATTTAAATTTACTGAATTTATTA 1046
Db 961 AACACAACTTTGGATTTTATATATTAACCTTGTGATTTAAATTTACTGAATTTATTA 1020
Qy 1047 GGGTGAATTTTGAAGTTGATTTCTCATATGACATAGTTCACTAAACCCGTGATGA 1106
Db 1021 GGGTGAATTTTGAAGTTGATTTCTCATATGACATAGTTCACTAAACCCGTGATGA 1080
Qy 1107 AAGTGAATTTATGTTCTCAGAACAAATGTCACAAAAAGACAAATATTTTTCACATGA 1166
Db 1081 AAGTGAATTTATGTTCTCAGAACAAATGTCACAAAAAGACAAATATTTTTCACATGA 1140
Qy 1167 ACCCTTGGCTGTAGTGGCTTTCTAGCTCACTCAAGCTCAAGCATCTTCAAGACGT 1226
Db 1141 ACCCTTGGCTGTAGTGGCTTTCTAGCTCACTCAAGCTCAAGCATCTTCAAGACGT 1200
Qy 1227 TTTCCCATATGCTGTATTTATTTTCTACATCACTCACTTCCCAATCATCTGGCT 1286
Db 1201 TTTCCCATATGCTGTATTTATTTTCTACATCACTCACTTCCCAATCATCTGGCT 1260
Qy 1287 GGCATCTCACAATGATGATGAAGCTGTCCTCCTCAAAACATCTGACATTTATTTGC 1346
Db 1261 GGCATCTCACAATGATGATGAAGCTGTCCTCCTCAAAACATCTGACATTTATTTGC 1320
Qy 1347 CAAATCAATACAAATCTTGAATTTTATATCGATCAATATTTTACAGTAATGATGATC 1406
Db 1321 CAAATCAATACAAATCTTGAATTTTATATCGATCAATATTTTACAGTAATGATGATC 1380
Qy 1407 AAACCTTCAATTTTAAACCTCTCTCTTGTGACAAACCTCTTAAAGAAATACAG 1466
Db 1381 AAACCTTCAATTTTAAACCTCTCTCTTGTGACAAACCTCTTAAAGAAATACAG 1440
Qy 1467 ATATATAGTAATATACCTCCACATCAGAGGTAGAACATCACTCTCCCTTGAGAT 1526
Db 1441 ATATATAGTAATATACCTCCACATCAGAGGTAGAACATCACTCTCCCTTGAGAT 1500
Qy 1527 CTTCACATAATCTAGTACTCACTTCCAAAGATGAGATGAGAAAGGAAACATAGTA 1586
Db 1501 CTTCACATAATCTAGTACTCACTTCCAAAGATGAGATGAGAAAGGAAACATAGTA 1560
Qy 1587 CTTTACAGGGGAAAAATGACAAATGACGCTTTCACCAAGTATCAAAATTAACGTAC 1646
Db 1561 CTTTACAGGGGAAAAATGACAAATGACGCTTTCACCAAGTATCAAAATTAACGTAC 1620
Qy 1647 CAGTGAATGATCAGATTTGTTCTAGATATCTTTCTAAAAATTCATATCCCAATC 1706
Db 1621 CAGTGAATGATCAGATTTGTTCTAGATATCTTTCTAAAAATTCATATCCCAATC 1680
Qy 1707 TAATTTAGGCTTAAACATCTCAGCAAACTCAAGTTGAAGCAATTTCAAAAATATCCT 1766
Db 1681 TAATTTAGGCTTAAACATCTCAGCAAACTCAAGTTGAAGCAATTTCAAAAATATCCT 1740
Qy 1767 GGGGTATTTAGATTTCTCTCAAACTGTAATAATCATGAAATTAAGGAATCTCGAG 1826
Db 1741 GGGGTATTTAGATTTCTCTCAAACTGTAATAATCATGAAATTAAGGAATCTCGAG 1800
Qy 1827 AAACAAATCAGACACATAGATAGAGAGATGAGACCAATGCAATGTCTTCTT 1886
Db 1801 AAACAAATCAGACACATAGATAGAGAGATGAGACCAATGCAATGTCTTCTT 1860
Qy 1887 GGATCGATCTTGAAACAGAAAAATGATCAAGTAATGAAAAATCTGATGAGTCTGAATGA 1946
Db 1861 GGATCGATCTTGAAACAGAAAAATGATCAAGTAATGAAAAATCTGATGAGTCTGAATGA 1920
Qy 1947 ATCTGAGATATTTTAAACATGATGATTTTATATCTTGAACAAATATAGAGGCTAA 2006
Db 1921 ATCTGAGATATTTTAAACATGATGATTTTATATCTTGAACAAATATAGAGGCTAA 1980

```

```

Qy 2007 TGTAGATGATTAAGCTTTAGAGAACTGAACCTGGGTGAGGGCTATAGCAATTTCTGT 2066
Db 1981 TGTAGATGATTAAGCTTTAGAGAACTGAACCTGGGTGAGGGCTATAGCAATTTCTGT 2040
Qy 2067 ACTATCTTACCAATTTTGGTAACTGTAAGAGCAATGCAAAATTAAGATATCTTGA 2126
Db 2041 ACTATCTTACCAATTTTGGTAACTGTAAGAGCAATGCAAAATTAAGATATCTTGA 2100
Qy 2127 AAAAAAAAAAAAAAAAAA 2144
Db 2101 AAAAAAAAAAAAAAAAAA 2118

RESULT 4
HSCRIP3G 2128 bp mRNA linear PRI 12-APR-1996
LOCUS HSCRIP3G
DEFINITION H.sapiens mRNA for cysteine-rich secretory protein-3.
ACCESSION X95240
VERSION X95240.1 GI:1262818
KEYWORDS CRISP-3 gene; cysteine-rich secretory protein-3.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 2128)
Kratzschmar J., Haendler B., Eberspaecher U., Roosterman D.,
Donner P. and Scheunling W.D.
The human cysteine-rich secretory protein (CRISP) family. Primary
structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3
Eur. J. Biochem. 236 (3), 827-836 (1996)
96270732
PUBMED 8665901
REFERENCE 2 (bases 1 to 2128)
AUTHORS Haendler B.
DIRECT SUBMISSION
TITLES Direct Submission
JOURNAL Submitted (18-JAN-1996) B. Haendler, Schering AG, ICMB, S109/517,
13342 Berlin, FRG
LOCATION/Qualifiers
FEATURES
SOURCE 1..2128
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1fb="testis cDNA library"
16..753
/gene="CRISP-3"
16..753
/gene="CRISP-3"
/codon_start=1
/product="cysteine-rich secretory protein-3"
/protein_id="CAA64527.1"
/db_xref="GI:1262819"
/db_xref="SWISS-PROT:P54108"
/translation="MTLPVLLPLFVAGLLSPFANEDKRPATLTLTOTVOREVN
KNEELRAVSPARNMLKEMKEAANAQKARNQNRHSPKDRMTSLKGENLYM
SSASSMSQALISWPEIDFDFGVGPRTPNVVGHYTVVYSSLVGCGNAYCPNO
KVLKLYKVCQYCPAGNMANRLVPEOGAPCASCPDNCDDGLCTNGCKEYEDLYSNKS
LKLTLTKYKQLVRDSCAKSCNCSNHY"
16..75
sig_peptide
/gene="CRISP-3"
BASE COUNT 2084..2089
ORIGIN 734 a 397 c 380 g 617 t
Query Match 98.6%; Score 2113.2; DB 9; Length 2128;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2115; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 27 CTGGAACCACTGCAATGATATTTCCAGTCTGTGTTCTGTTGCTGGGCTGCTT 86
Db 1 CTGGAACCACTGCAATGATATTTCCAGTCTGTGTTCTGTTGCTGGGCTGCTT 60
Qy 87 CCATCTTTCCAGCAAAATGAAGATCCGCTTTTACTCTTTGTTAACCCACCA 146

```

Db 61 GCATCTTTCCAGCAATGATAGATAGATCCCGCTTTTACTGCTTTGTTAACCCACCA 120
 Oy 147 ACACAAGTCCAAAGGAGATGTGTAATAGACACAATGAAGAGAGACAGATATCTCC 206
 Db 121 ACACAAGTCCAAAGGAGATGTGTAATAGACACAATGAAGAGAGACAGATATCTCC 180
 Oy 207 CCTGGAGAAATGCTGTAAGATGGAATGGAACCAAGAGAGCTGCACCAATGCCCAAG 266
 Db 181 CCTGGAGAAATGCTGTAAGATGGAATGGAACCAAGAGAGCTGCACCAATGCCCAAG 240
 Oy 267 TGGGCAAAACAGTGCATATACAGACAGACAGTAAACCAAGAGATGCAATGCAAGTCAAA 326
 Db 241 TGGGCAAAACAGTGCATATACAGACAGTAAACCAAGAGATGCAATGCAAGTCAAA 300
 Oy 327 TGTGTGAGAAATCTCTACATGTCAAGTGCACCAAGCTATGTCACAAAGCTCAAGC 386
 Db 301 TGTGTGAGAAATCTCTACATGTCAAGTGCACCAAGCTATGTCACAAAGCTCAAGC 360
 Oy 387 TGTGTGAGAAATCTCTACATGTCAAGTGCACCAAGCTATGTCACAAAGCTCAAGC 446
 Db 361 TGTGTGAGAAATCTCTACATGTCAAGTGCACCAAGCTATGTCACAAAGCTCAAGC 420
 Oy 447 GTTGACATTTATACAGAGTGTGTGTGCTCTCTACCTGCTGAGATGGAATATGCG 506
 Db 421 GTTGACATTTATACAGAGTGTGTGTGCTCTCTACCTGCTGAGATGGAATATGCG 480
 Oy 507 TACTGTCCCAATCAAAAGTCTTAAATATCTACTATGTTGCCAATATGCTGCTGCT 566
 Db 481 TACTGTCCCAATCAAAAGTCTTAAATATCTACTATGTTGCCAATATGCTGCTGCTG 540
 Oy 567 AATTGGGCAATAGACTATATGTCCTTATGAACAAGAGACACTGTTGCCAGTTGCCCA 626
 Db 541 AATTGGGCAATAGACTATATGTCCTTATGAACAAGAGACACTGTTGCCAGTTGCCCA 600
 Oy 627 GATACTGTAGAGATGAGATGACACCAATGTTGCAATGAGAGATCTCTATAGTAC 686
 Db 601 GATACTGTAGAGATGAGATGACACCAATGTTGCAATGAGAGATCTCTATAGTAC 660
 Oy 687 TGTAAAAGTTTAAAGCTCATTAACTCTGTAACATCACTGCTGACGAGAGTTGCAAG 746
 Db 661 TGTAAAAGTTTAAAGCTCATTAACTCTGTAACATCACTGCTGACGAGAGTTGCAAG 720
 Oy 747 GCATCTGCAATTTGTTCAACAGCAATTTATTAATACGATTAACACAGGAGTGGCTA 806
 Db 721 GCATCTGCAATTTGTTCAACAGCAATTTATTAATACGATTAACACAGGAGTGGCTA 780
 Oy 807 TGTAGAGAGAGTCAAGATTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 866
 Db 781 TGTAGAGAGAGTCAAGATTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 840
 Oy 867 AGCTAGAGAAATGTAGGCAATGTTGATACACATTTGATTTCAAAATGTTTCTCTGAT 926
 Db 841 AGCTAGAGAAATGTAGGCAATGTTGATACACATTTGATTTCAAAATGTTTCTCTGAT 900
 Oy 927 CTGCTTTTATTTTCAAAAATTTTTCATACAAATGTTTAAAGAGAAACAAATCTAT 986
 Db 901 CTGCTTTTATTTTCAAAAATTTTTCATACAAATGTTTAAAGAGAAACAAATCTAT 960
 Oy 987 AACACAACATTTGATTTTATATATAACTTTGATTTAAATTTACTGAATTTATTA 1046
 Db 961 AACACAACATTTGATTTTATATATAACTTTGATTTAAATTTACTGAATTTATTA 1020
 Oy 1047 GGGTGAATAATTTTGAAGTGTATCTCTATGACTAAGTCACTAAACCTGGATTGA 1106
 Db 1021 GGGTGAATAATTTTGAAGTGTATCTCTATGACTAAGTCACTAAACCTGGATTGA 1080
 Oy 1107 AAGTGAATAATTTTCTCTAGAACAAATGTACAAAAGAACAAATATTTTTCACATGA 1166
 Db 1081 AAGTGAATAATTTTCTCTAGAACAAATGTACAAAAGAACAAATATTTTTCACATGA 1140
 Oy 1167 ACCCTTGGCTGATGTTGCTTCTAGCTCACTTAAGGCTTAAGCATCTTAAGAGCT 1226
 Db 1141 ACCCTTGGCTGATGTTGCTTCTAGCTCACTTAAGGCTTAAGCATCTTAAGAGCT 1200

Oy 1227 TTTCCATATGCTGTCTTAATTTCTTTTACATCACTTACACCTTCTTCCCAATCATGCT 1286
 Db 1201 TTTCCATATGCTGTCTTAATTTCTTTTACATCACTTACACCTTCTTCCCAATCATGCT 1260
 Oy 1287 GGCATCTCACAATTTGATGTAAGCTGTCTCTCTTAAACCAATCCGACTTTTATTTTGC 1346
 Db 1261 GGCATCTCACAATTTGATGTAAGCTGTCTCTCTTAAACCAATCCGACTTTTATTTTGC 1320
 Oy 1347 CAAATCAATACATCTCTTGAATTTTATCTGCTAATTTTACAGTAGAATATGATC 1406
 Db 1321 CAAATCAATACATCTCTTGAATTTTATCTGCTAATTTTACAGTAGAATATGATC 1380
 Oy 1407 AAACCTTCAATTTTAAACCTCTCTCTTCTTGAACAAACCTCTTAAACCAATCAAG 1466
 Db 1381 AAACCTTCAATTTTAAACCTCTCTCTCTTCTTGAACAAACCTCTTAAACCAATCAAG 1440
 Oy 1467 ATATATATAGTAAATATACCTTCACTCAAGAGTAGAATCTCACTCTCTCTGAGT 1526
 Db 1441 ATATATATAGTAAATATACCTTCACTCAAGAGTAGAATCTCACTCTCTCTGAGT 1500
 Oy 1527 CTTCATTAATCAGTCACTCACTTCCAAAGAGTGTGATGGAAGGAAACATAGTAA 1586
 Db 1501 CTTCATTAATCAGTCACTCACTTCCAAAGAGTGTGATGGAAGGAAACATAGTAA 1560
 Oy 1587 CTTTACAGGAGGAAACCAATGACAAATGACGCTTCCACCAAGTATCAAAATTAAGCTAC 1646
 Db 1561 CTTTACAGGAGGAAACCAATGACAAATGACGCTTCCACCAAGTATCAAAATTAAGCTAC 1620
 Oy 1647 CAGTGTATGATTTATCAATTTGTTCTAGATTAATCTTTTAAACCAATTTATCCATC 1706
 Db 1621 CAGTGTATGATTTATCAATTTGTTCTAGATTAATCTTTTAAACCAATTTATCCATC 1680
 Oy 1707 TAATATGAGTAAACCAATCAGCAACCAATCAAGTTAAAGAGATCTTACAAATATCCCT 1766
 Db 1681 TAATATGAGTAAACCAATCAGCAACCAATCAAGTTAAAGAGATCTTACAAATATCCCT 1740
 Oy 1767 GGGGTATTTTATGATTTCTCTCAAAACCTGTAAACATCATGAAATTAAGGAATCCCTGAG 1826
 Db 1741 GGGGTATTTTATGATTTCTCTCAAAACCTGTAAACATCATGAAATTAAGGAATCCCTGAG 1800
 Oy 1827 AAACATCAGACCAACATGACATGAGAGACATGAGAGAGTGAAGTGAATGCTTCT 1886
 Db 1801 AAACATCAGACCAACATGACATGAGAGACATGAGAGAGTGAAGTGAATGCTTCT 1860
 Oy 1887 GGATCAGATCTGGAACAGAAACCAATGATGATGATGATGATGATGATGATGATGATGAT 1946
 Db 1861 GGATCAGATCTGGAACAGAAACCAATGATGATGATGATGATGATGATGATGATGATGAT 1920
 Oy 1947 ATCTGAGATTTTATTAACATGATGATGATGATGATGATGATGATGATGATGATGAT 2006
 Db 1921 ATCTGAGATTTTATTAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
 Oy 2007 TGTAGATGATTAAGTGTAGAGAACTGAACAGTGGGAGGAGTATCTGGAATCTGCT 2066
 Db 1981 TGTAGATGATTAAGTGTAGAGAACTGAACAGTGGGAGGAGTATCTGGAATCTGCT 2040
 Oy 2067 ACTATCTTACCAATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2126
 Db 2041 ACTATCTTACCAATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
 Oy 2127 AAAAAAAAAAAAAAAAAA 2144
 Db 2101 AAAAAAAAAAAAAAAAAA 2118

RESULT 5
 AC010779/c 151752 bp DNA linear HTG 22-JUL-2000
 LOCUS AC010779 Homo sapiens clone Rpl1-34d, WORKING DRAFT SEQUENCE, 11 unordered
 DEFINITION pieces.
 ACCESSION AC010779
 VERSION AC010779.4 GI:9369464

KEYWORDS
SOURCE
ORGANISM

HTG: HTGS_PHASE1; HTGS_DRAFT.

Homo sapiens.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.REFERENCE
AUTHORS

1 (bases 1 to 151752)

Birten, B., Linton, L., Nusbaum, C. and Lander, E.

Unpublished

TITLE
JOURNAL
REFERENCE
AUTHORS

2 (bases 1 to 151752)

Birten, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Bouknight, B.,
Brown, A., Castle, A., Collangelo, M., Collins, S., Collamore, A.,
Coke, P., Dearlano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., MacDonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Tittel, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (22-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 22, 2000 this sequence version replaced g1:7321560.All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L2699

Center clone name: 3_A_4

Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 142917 bases at least Q40

Consensus quality: 148664 bases at least Q20

Insert size: 150000; agarose-fp

Quality coverage: 4.8 in Q20 bases; sum-of-contigs

Quality coverage: 4.8 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
 consists of 11 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 1901: contig of 1901 bp in length
 * 1902 2001: gap of 100 bp
 * 2002 2001: contig of 2630 bp in length
 * 4632 4731: gap of 100 bp
 * 4732 9746: contig of 5015 bp in length
 * 9747 9846: gap of 100 bp
 * 9847 17669: contig of 7823 bp in length
 * 17670 17769: gap of 100 bp
 * 17770 29964: contig of 12195 bp in length
 * 29965 30064: gap of 100 bp
 * 30065 43497: contig of 13433 bp in length
 * 43498 43597: gap of 100 bp
 * 43598 60340: contig of 16743 bp in length
 * 60341 60440: gap of 100 bp

60441 75774: contig of 15334 bp in length
 * 75775 75874: gap of 100 bp
 * 75875 93175: contig of 17301 bp in length
 * 93176 93275: gap of 100 bp
 * 93276 119039: contig of 25764 bp in length
 * 119040 119139: gap of 100 bp
 * 119140 151752: contig of 32613 bp in length.

FEATURES

source

1. 151752

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_Rp11-3A4"

/clone_11b="RP11-11 Human Male BAC"

1. 1901

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

BASE COUNT 49163 a 27268 c 27224 g 47065 t 1032 others

ORIGIN

Query Match

Best Local Similarity 99.8%; Score 1485.2; DB 2; Length 151752;

Matches 1487; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

652 CCAATGGTGAAGTGAAGATCTCTATAGTACTGTAAAGTTGAAGCTCACAATTAA 711
 Db 80106 CCAATGGTGAAGTGAAGATCTCTATAGTACTGTAAAGTTGAAGCTCACAATTAA 80047
 QY 712 CCTGTAACATGATGGTGCAGGAGACAGTTCAGCAAGCATCTGCATTTGCAACAGCA 771
 Db 80046 CCTGTAACATGATGGTGCAGGAGACAGTTCAGCAAGCATCTGCATTTGCAACAGCA 79987
 QY 772 TTTATTAATAGCATTTACACACCGAGTAGGCGTATGTAGAGAGAGTACATTATCTAC 831
 Db 79986 TTTATTAATAGCATTTACACACCGAGTAGGCGTATGTAGAGAGAGTACATTATCTAC 79927
 QY 832 TAGATTTGGCATCTACTAGATTTAATATATACATATACAGTGAAGTGAAGCATGTTTG 891
 Db 79926 TAGATTTGGCATCTACTAGATTTAATATATACATATACAGTGAAGTGAAGCATGTTTG 79867
 QY 892 ATACATTTGATTTAAATGTTTCTCTGATGCTGTTTATTTATTTACAAAATATAT 951
 Db 79866 ATACATTTGATTTAAATGTTTCTCTGATGCTGTTTATTTATTTACAAAATATAT 79807
 QY 952 TTTCATACAAATGTTTAAAGAAACAAATATCTATACAAACATTTGATTTATATA 1011
 Db 79806 TTTCATACAAATGTTTAAAGAAACAAATATCTATACAAACATTTGATTTATATA 79747
 QY 1012 TAAACTTTGATTTAAATTTACTGAATTTAATTTAGGTGAAATTTTGAAGCTTATAT 1071
 Db 79746 TAAACTTTGATTTAAATTTACTGAATTTAATTTAGGTGAAATTTTGAAGCTTATAT 79687

QY 1072 CTCATATGACTAAGTTCACATAAAACCTGATGTAAGTGAATAATATGTTCTAGAACAA 1131
 |||||||
 Db 79686 CTCATATGACTAAGTTCACATAAAACCTGATGTAAGTGAATAATATGTTCTAGAACAA 79627
 |||||||
 QY 1132 AATGTACAAAAGACAAATATATTTTACATGAGACCTTGCTGTAGTTGCTTTCT 1191
 |||||||
 Db 79626 AATGTACAAAAGACAAATATATTTTACATGAGACCTTGCTGTAGTTGCTTTCT 79567
 |||||||
 QY 1192 AGCTCCACTAAGGCTAAGATGCTCAAGACGTTTCCATATGCTGTATATTTCT 1251
 |||||||
 Db 79566 AGCTCCACTAAGGCTAAGATGCTCAAGACGTTTCCATATGCTGTATATTTCT 79567
 |||||||
 QY 1252 TTCACATTCACCCCTTCTCCCAATCATCTGCTGGCATCTCCCAATGAGTTGAAGC 1311
 |||||||
 Db 79506 TTCACATTCACCCCTTCTCCCAATCATCTGCTGGCATCTCCCAATGAGTTGAAGC 79447
 |||||||
 QY 1312 TGTTCCTCTTAAACATCTGACTTTTATTTTGGCAAAATCATACATCTCTGAAT 1371
 |||||||
 Db 79446 TGTTCCTCTTAAACATCTGACTTTTATTTTGGCAAAATCATACATCTCTGAAT 79387
 |||||||
 QY 1372 TTTATCTGCATAAATTTTACAGTATGATCAAACTTCATTTTAAACCTCTCT 1431
 |||||||
 Db 79386 TTTATCTGCATAAATTTTACAGTATGATCAAACTTCATTTTAAACCTCTCT 79327
 |||||||
 QY 1432 CTCTTGCACAAACTCTCTTAAAAAGAAATCAAGATATATAGTAAATACCTCTCACT 1491
 |||||||
 Db 79326 CTCTTGCACAAACTCTCTTAAAAAGAAATCAAGATATATAGTAAATACCTCTCACT 79267
 |||||||
 QY 1492 CAAGGAGTAACTCATCTCTCTCTGAGTCTTCACTAAATCAGTACCTCACT 1551
 |||||||
 Db 79266 CAAGGAGTAACTCATCTCTCTCTGAGTCTTCACTAAATCAGTACCTCACT 79207
 |||||||
 QY 1552 CCAAGAGTGAATGGAAGGGAAGGGAACATAGTAATTTACAGGGGAGAAAATGACAAA 1611
 |||||||
 Db 79206 CCAAGAGTGAATGGAAGGGAAGGGAACATAGTAATTTACAGGGGAGAAAATGACAAA 79147
 |||||||
 QY 1612 TGAGCTCTTCCACCAAGTATCAAAATTAAGCTCACCAGTATAGTATTCAGATTGTT 1671
 |||||||
 Db 79146 TGAGCTCTTCCACCAAGTATCAAAATTAAGCTCACCAGTATAGTATTCAGATTGTT 79087
 |||||||
 QY 1672 CTAGATATCTTTCTAAAAATTCATATCCCAATCTATATGAGTAAACATCCACCA 1731
 |||||||
 Db 79086 CTAGATATCTTTCTAAAAATTCATATCCCAATCTATATGAGTAAACATCCACCA 79027
 |||||||
 QY 1732 AACTCAAGTGAAGACATTTCTACAAATATCCCTGGGATATTTAGAGTATTCCTCAAA 1791
 |||||||
 Db 79026 AACTCAAGTGAAGACATTTCTACAAATATCCCTGGGATATTTAGAGTATTCCTCAAA 78967
 |||||||
 QY 1792 ACTGTAAAAATCATGAGAAATAGGGAATCCTGAGAAAATCAGACACATGAGACT 1851
 |||||||
 Db 78966 ACTGTAAAAATCATGAGAAATAGGGAATCCTGAGAAAATCAGACACATGAGACT 78907
 |||||||
 QY 1852 AAGGAGCATGAGGCCAAATGCAATGCTCTTGATGATGATCCCTGGAACAGAAAAG 1911
 |||||||
 Db 78906 AAGGAGCATGAGGCCAAATGCAATGCTCTTGATGATGATCCCTGGAACAGAAAAG 78847
 |||||||
 QY 1912 ATCAATAATGAAAAAATGATGAAGTCTGAATAGATCTGAGATATTTTAAACAGTAGTG 1971
 |||||||
 Db 78846 ATCAATAATGAAAAAATGATGAAGTCTGAATAGATCTGAGATATTTTAAACAGTAGTG 78787
 |||||||
 QY 1972 TTGATTTTCTTAATCTTGCAAAATATAGCAGGGTATGTAAGTGAACGTTAGAGAAC 2031
 |||||||
 Db 78786 TTGATTTTCTTAATCTTGCAAAATATAGCAGGGTATGTAAGTGAACGTTAGAGAAC 78727
 |||||||
 QY 2032 TGAACCTGGGTGAGGCGATCTAGGAATCTCTGACTATCTTACCAATTTTGGTAAAG 2091
 |||||||
 Db 78726 TGAACCTGGGTGAGGCGATCTAGGAATCTCTGACTATCTTACCAATTTTGGTAAAG 78667
 |||||||
 QY 2092 TCTAAGAAGCAATGCAAAATTAAGATATCTTGAAGAAAAAATTAAGAA 2141
 |||||||
 Db 78666 TCTAAGAAGCAATGCAAAATTAAGATATCTTGAAGAAAAAATTAAGAA 78617
 |||||||

RESULT 6

HSDJ442L6/c

LOCUS

HSDJ442L6

DEFINITION

Human DNA sequence from clone RP3-442L6 on chromosome 6. Contains the TPX1 gene encoding a Testis specific protein 1 (CRISP-2, GAPD5), the 5' end of RHAG gene encoding Rhesus blood group-associated glycoprotein (RH50a), the 3' end of SPG28 encoding a cysteine-rich secretory protein-3 (CRISP-3), ESTs, STSS and GSSs, complete sequence.

ACCESSION

AL121950

VERSION

AL121950.8

KEYWORDS

GT:9581779

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Parker, A.

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (29-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 On Jul 28, 2000 this sequence version replaced gi:9408740.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; SW, SWISSPROT; Tr, TrEMBL; Wp, WormPeP; Information on the WormPeP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone/corliss of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping group. Further information can be found at <http://www.sanger.ac.uk/RGP/Chr6>
 RP3-442L6 is from the library RP3-3 constructed at the Rosewell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>
 VECTOR: pCIPAC2
 This sequence is the entire insert of clone RP3-442L6 The true left end of clone RP3-417L20 is at 85775 in this sequence.

FEATURES

Location/Qualifiers

1..118524

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="6"

/clone="RP3-442L6"

/clone_id="RP3-3"

/complement(1..25721)

/gene="RHAG"

complement(join(1..27,1279..1416,3569..3735,4506..4653,6950..7100,8061..8244,25538..25721))

/product="dj442L6.1 (Rhesus blood group-associated glycoprotein (RH50a))"

/note="match: CDNAS: Em:X6594 Em:AF179682 Em:AF031548 Em:AF187847 Em:AF178841 Em:AF058917 Em:AF179684 Em:AB015467"

/evidence="not_experimental"

complement(join(1..27,1279..1416,3569..3735,4506..4653,6950..7100,8061..8244,25538..25721))

/gene="RHAG"

/note="Continues in clone dj417L20 (Em:AL121974)"

CDS


```
match: proteins: Tr:09V373 Tr:088300 Tr:07484 Tr:09UBD6
Tr:09QXP0"
/codon_start=1
/evidence=not_experimental
/product="J4442L6.1 (Rhesus blood group-associated
glycoprotein (RH50A))"
/protein_id="CAC10519.1"
/db_xref="GI:10443472"
/translation="MRTFPLMAIVLEIAMIVLGLPEYETDQVLEQLNITKPTDM
GIEFELPLFDVHVMIFVGFPLMTFLKATGSSVGINLLVAALGLQWGTIVQGLQ
SQGKFNIGIKMINADFSAAATVLISFGAVLAKTSPQMLIMTILIEVFANIELVLS
EIPRASDIGAMTTHAFGAYGLAVAGILYSGRKHGENESAYSDLPANIGLFL
MMPSPNSAIAEPDGKOCRAIVNTYPSLACVLTAFFSSIVEHGRKLVNHIQNAI
LAGVAVGTCAADMAHPGSMIIGSINAGMSVLGKFLPTTLRLI"
repeat_region
257..446
/notes="L1M47 repeat: matches 5986..6165 of consensus"
623..930
/notes="AluX repeat: matches 1..301 of consensus"
1482..1913
/notes="match: GSS: Em:AQ813163"
1952..2249
/notes="AluX repeat: matches 3..300 of consensus"
2602..2746
/notes="L1MD3 repeat: matches 7582..7740 of consensus"
2785..2832
/notes="L1MC3 repeat: matches 7478..7525 of consensus"
3064..3383
/notes="AluSq repeat: matches 1..311 of consensus"
3823..4131
/notes="AluSg repeat: matches 1..305 of consensus"
4288..4406
/notes="FLAM_A repeat: matches 1..122 of consensus"
4820..5005
/notes="MIR repeat: matches 20..211 of consensus"
5084..5257
/notes="FRAM repeat: matches 0..172 of consensus"
5951..6387
/notes="L2 repeat: matches 2246..2698 of consensus"
6673..6828
/notes="MIR repeat: matches 95..262 of consensus"
complement(7407..8012)
/genes="RHAG"
/notes="match: GSS: Em:B87311"
7685..8049
/notes="MLT1A1 repeat: matches 1..365 of consensus"
8340..8442
/notes="MIR repeat: matches 4..106 of consensus"
9022..9137
/notes="MIR repeat: matches 49..172 of consensus"
9432..9712
/notes="AluX repeat: matches 1..262 of consensus"
10074..10388
/notes="LTR16A repeat: matches 14..381 of consensus"
10815..10880
/notes="33 copies 2 mer tt 66% conserved"
10939..11098
/notes="THE1B repeat: matches 1..167 of consensus"
11099..11389
/notes="AluSg repeat: matches 2..297 of consensus"
11390..11609
/notes="THE1B repeat: matches 167..364 of consensus"
11713..12019
/notes="AluY repeat: matches 1..304 of consensus"
12444..12662
/notes="MER20 repeat: matches 1..217 of consensus"
13096..13385
/notes="AluX repeat: matches 2..290 of consensus"
13474..13553
/notes="MIR repeat: matches 74..182 of consensus"
complement(13493..13920)
/genes="RHAG"
/notes="match: GSS: Em:AQ813968"
13541..13606
repeat_region

misc_feature
13971..14526
/notes="L2 repeat: matches 2637..2707 of consensus"
14139..14391
/notes="match: GSS: Em:AQ539951"
14620..14983
/notes="MIR repeat: matches 5..262 of consensus"
14984..16480
/notes="THE1B repeat: matches 1..364 of consensus"
/notes="THE1B-INTERNAL repeat: matches 1..1580 of
consensus"
repeat_region
16481..16840
/notes="THE1B repeat: matches 1..364 of consensus"
17057..17296
/notes="AluJo repeat: matches 31..282 of consensus"
17297..18994
/notes="P1R5 repeat: matches 743..1686 of consensus"
19639..19831
/notes="L1MD3 repeat: matches 7448..7634 of consensus"
19863..20630
/notes="L1MC4 repeat: matches 6630..7474 of consensus"
20891..21016
/notes="Charlie1a repeat: matches 1..128 of consensus"
21013..21609
/notes="Charlie1a repeat: matches 331..933 of consensus"
22116..22212
/notes="L2 repeat: matches 2589..2695 of consensus"
22790..22844
/notes="L2 repeat: matches 2617..2671 of consensus"
22860..23218
/notes="L1M44 repeat: matches 5950..6299 of consensus"
23216..23511
/notes="L1M2 repeat: matches 5508..5806 of consensus"
23683..23832
/notes="L1M43 repeat: matches 6152..6302 of consensus"
23839..23880
/notes="21 copies 2 mer tg 88% conserved"
24216..24538
/notes="match: GSS: Em:AQ071395"
27538..27945
/notes="L2 repeat: matches 2356..2750 of consensus"
28904..29192
/notes="AluX repeat: matches 1..289 of consensus"
29656..30138
/notes="Charlie4a repeat: matches 1..507 of consensus"
30286..30707
/notes="match: GSS: Em:AQ465695"
30451..30583
/notes="MIR repeat: matches 1..134 of consensus"
30513..30810
/notes="match: GSS: Em:AQ114043"
30770..31070
/notes="AluY repeat: matches 1..303 of consensus"
31388..31459
/notes="36 copies 2 mer ta 69% conserved"
31808..32127
/notes="AluSc repeat: matches 1..309 of consensus"
32391..32514
repeat_region

Query Match 69.0%; Score 1480.4; DB 9; Length 118524;
Best Local Similarity 99.6%; Pred. No. 4e-263;
Matches 1484; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 652 CCAATGCTTCAAGTACAGATCTATAGTAACTTAAAGTTAACTACATTA 711
|||||
DB 117739 CCAATGCTTCAAGTACAGATCTATAGTAACTTAAAGTTAACTACATTA 117680

QY 712 CCTGTAACATCAGTTGGTAGGAGACAGTGCAGGACCTCCTGCAATTGTTCAACAGCA 771
|||||
DB 117679 CCTGTAACATCAGTTGGTAGGAGACAGTGCAGGACCTCCTGCAATTGTTCAACAGCA 117620

QY 772 TTTATTAATACGATTACACACGAGTACAGGCTATGTAGAGAGACTAGATTATCTAC 831
|||||
DB 117619 TTTATTAATACGATTACACACGAGTACAGGCTATGTAGAGAGACTAGATTATCTAC 117560
```

QY 832 TTAGATTGGCATCTACTAGATTACATATAGTGAAGAAATTTAGGCATGTTG 891
 DB 117559 TTAGATTGGCATCTACTAGATTACATATAGTGAAGAAATTTAGGCATGTTG 117500
 QY 892 ATACACATTTGATTTCAAAATGTTTTCTCTGATCTGTTTTATTTTACAAAATAT 951
 DB 117499 ATACACATTTGATTTCAAAATGTTTTCTCTGATCTGTTTTATTTTACAAAATAT 117440
 QY 952 TTTCATACAAATGTTAAAAAGAACAAATCTATACAAACATTTGATTTTATATA 1011
 DB 117439 TTTCATACAAATGTTAAAAAGAACAAATCTATACAAACATTTGATTTTATATA 117380
 QY 1012 TAAACCTTGATTTAAATTTACTGAAATTAATTTAGGGGAAATTTTAAAGTTGAT 1071
 DB 117379 TAAACCTTGATTTAAATTTACTGAAATTAATTTAGGGGAAATTTTAAAGTTGAT 117320
 QY 1072 CTCATATGACATAGTTCACTAAACCCCTGATTTGAAGTGAATAATTTGCTCTAGACA 1131
 DB 117319 CTCATATGACATAGTTCACTAAACCCCTGATTTGAAGTGAATAATTTGCTCTAGACA 117260
 QY 1132 AATGTACAAAAGAACATATATTTTACATGAAACCTTGCTGCTAGTGGCTTCT 1191
 DB 117259 AATGTACAAAAGAACATATATTTTACATGAAACCTTGCTGCTAGTGGCTTCT 117200
 QY 1192 AGCTCCACTCTAAGCTAAGCATCTTCAAGAGAGTTTCCCATATGCTGCTTAAATCT 1251
 DB 117199 AGCTCCACTCTAAGCTAAGCATCTTCAAGAGAGTTTCCCATATGCTGCTTAAATCT 117140
 QY 1252 TTCATCATCTACCCCTCTTCCCAATCATCTGGCTGGCATCTCACAATTTGAGTTGAGC 1311
 DB 117139 TTCATCATCTACCCCTCTTCCCAATCATCTGGCTGGCATCTCACAATTTGAGTTGAGC 117080
 QY 1312 TGTCTCTCTTAAACAAATCTGACTTTTATTTGGCAAAATCAATATCATCTCTTGAAT 1371
 DB 117079 TGTCTCTCTTAAACAAATCTGACTTTTATTTGGCAAAATCAATATCATCTCTTGAAT 117020
 QY 1372 TTTTATCTGCATTAATTTTACAGTAATATGATCAAACTTATTTTAAACCTCTCT 1431
 DB 117019 TTTTATCTGCATTAATTTTACAGTAATATGATCAAACTTATTTTAAACCTCTCT 116960
 QY 1432 CTCCTTGACAAAATCTCTCTTAAAAAAGAAATACAGATATATATAGTAATATACCTCCACT 1491
 DB 116959 CTCCTTGACAAAATCTCTCTTAAAAAAGAAATACAGATATATATAGTAATATACCTCCACT 116900
 QY 1492 CAAGGAGTAGAAGTCACTCCCTCTCTCTGATGCTTCACTTAAATCAGTCACTCACT 1551
 DB 116899 CAAGGAGTAGAAGTCACTCCCTCTCTCTGATGCTTCACTTAAATCAGTCACTCACT 116840
 QY 1552 CCAGAGAGTGAAGTATGGAAGGAACAATAGTAACTTACAGGGGAGAAATATGACAAA 1611
 DB 116839 CCAGAGAGTGAAGTATGGAAGGAACAATAGTAACTTACAGGGGAGAAATATGACAAA 116780
 QY 1612 TGACCTCTTACCAAGATGATCAAAATTAAGCTCACAGTATAGTCACTTCACTTTGTT 1671
 DB 116779 TGACCTCTTACCAAGATGATCAAAATTAAGCTCACAGTATAGTCACTTCACTTTGTT 116720
 QY 1672 CTAGATATCTTTCTTAAATTTCAATATCCCAATCTATATATAGCTTAAATATCCAGCA 1731
 DB 116719 CTAGATATCTTTCTTAAATTTCAATATCCCAATCTATATATAGCTTAAATATCCAGCA 116660
 QY 1732 AACTCAAGTGAAGACATTTCTACAAAATATCCCTGGGATTTTAAAGTATTTCTCAAA 1791
 DB 116659 AACTCAAGTGAAGACATTTCTACAAAATATCCCTGGGATTTTAAAGTATTTCTCAAA 116600
 QY 1792 ACTGTAAATATCAGGAAATATAGGAAATCTGAGAAACATACAGACACATAGACT 1851
 DB 116599 ACTGTAAATATCAGGAAATATAGGAAATCTGAGAAACATACAGACACATAGACT 116540
 QY 1852 AAGGAGACATGTGAGCCAAATGCAATGTGCTTTTGATCAGATCTTGGACAGAGAAAAG 1911
 DB 116539 AAGGAGACATGTGAGCCAAATGCAATGTGCTTTTGATCAGATCTTGGACAGAGAAAAG 116480

QY 1912 ATCAGTATGAAATCTATGACGTCTGATATGAAATGAAATTTTAAACAGTACTG 1971
 DB 116479 ATCAGTATGAAATCTATGACGTCTGATATGAAATGAAATTTTAAACAGTACTG 116420
 QY 1972 TTGATTTCTTAATCTTGAACAAATATFAGCAGGGTAAATGATGATTAAGTGAAGAAC 2031
 DB 116419 TTGATTTCTTAATCTTGAACAAATATFAGCAGGGTAAATGATGATTAAGTGAAGAAC 116360
 QY 2032 TGAAGTGGGAGGAGCTATCTGATGAAATTTCTGATATCTTACCAATTTTGGGTAG 2091
 DB 116359 TGAAGTGGGAGGAGCTATCTGATGAAATTTCTGATGAAATTTTGGGTAG 116300
 QY 2092 TCTAGAAAGCAATGCAAAATATGAAATTTCTTGAAGAAAAAATTTTGGGTAG 2141
 DB 116299 TCTAGAAAGCAATGCAAAATATGAAATTTCTTGAAGAAAAAATTTTGGGTAG 116250

RESULT 7
 HUMTPX1A
 LOCUS HUMTPX1A 1335 bp mRNA linear PRI 14-JAN-1995
 DEFINITION Human testis-specific protein (Tpx-1) mRNA, complete cds.
 ACCESSION M25532 J04741
 VERSION M25532.1 GI:339882
 KEYWORDS testis-specific protein.
 SOURCE Human adult testis, cDNA to mRNA, clones H4-1 and p3-1.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1335)
 Kashiwara, M., Gutknecht, J., Brew, K., Spurr, N. and Goodfellow, P. N.
 Cloning and mapping of a testis-specific gene with sequence similarity to a sperm-coating glycoprotein gene
 Genomics 5 (3), 527-534 (1989)
 JOURNAL MEDLINE 90129048
 PUBMED 2613236
 COMMENT Draft entry and computer-readable sequence for [1] kindly submitted by M. Kashiwara, 09-JUN-1989.
 FEATURES
 source location/Qualifiers
 1..1335 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="6p21-qter"
 1..1335 /gene="GAPDL5"
 <1..1335 /gene="GAPDL5"
 /product="testis-specific protein mRNA"
 190..921 /gene="GAPDL5"
 /note="testis-specific protein precursor"
 /codon_start=1 /protein_id="AA61220.1"
 /db_xref="GI:339883"
 /db_xref="GDB:G00-119-257"
 /translation="MALLPVLELVLTLPSPARCKDPAFTALLTOLQVREYVNH NELKRAVSPPRSNNLKHESREVTNNMORANRCKTLDSHPEDRKSTRGENLIMSS DPISWSSAIOSWDEILDFYVGVGPKSPNAVGHYTLQVSTYOVCGIAYCPNDS LKYYVQOYCPAGNNMNRKNTPOOGGPPACBDCKGLCTNCSQYODLISNCSLK NTACSEHLELKEKCAICLENKIY"
 190..249 /gene="GAPDL5"
 /note="testis-specific signal peptide (put.): putative"
 mat_peptide 250..918 /gene="GAPDL5"
 /product="testis-specific protein (put.): putative"
 BASE COUNT 428 a 279 c 271 g 355 t 2 others
 ORIGIN Chromosome 6.
 Query Match 24.0%; Score 515.2; DB 9; Length 1335;
 Best Local Similarity 68.8%; Pred. No. 4.4e-85;
 Matches 798; Conservative 0; Mismatches 323; Indels 39; Gaps 5;
 QY 1 TGATGAACAATACTTACTCTGCTGGAACCA---CTGCATGACATTTATCCCA 56

Db 171 TAATAAGATAGATATTTCATCTGCTCAGAAACCAACATTTCCAGCATGGCTTTACTA 230
 Oy 57 GTGCTGTGTTCTCTGCTGCTGGCTGCTTCCATCTTTTCCAGCAATGAGATAGAT 116
 Db 231 CCGGCTGTGTTCTCTGCTGCTGGCTGCTTCCATCTTTTCCAGCAATGAGAT 287
 Oy 117 CCGGCTGTGTTCTCTGCTGCTGGCTGCTTCCATCTTTTCCAGCAATGAGAT 176
 Db 288 CCGGCTGTGTTCTCTGCTGCTGGCTGCTTCCATCTTTTCCAGCAATGAGAT 347
 Oy 177 CACATGAGATGAGAGAGAGATATCTCCCTGCGAGAACATCTGAGATGAGATG 236
 Db 348 CACATGAGATGAGAGAGAGATATCTCCCTGCGAGAACATCTGAGATGAGATG 407
 Oy 237 AACAAAGAGCTGACGAAATGCCAAAGTGGCAACCAAGTGCATTTACAGACAGT 296
 Db 408 AGCAGAGAGCTAACAGATGCCAAAGTGGCAACCAAGTGCATTTACAGACAGT 467
 Oy 297 AACCAAGGATGAGATGCTAAATGTTGGTGAATCTGATCTGATGCTGCTGCTG 356
 Db 468 GATCCAGAGAGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 527
 Oy 357 CCGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 416
 Db 528 CCGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 587
 Oy 417 GGTGAGGAGAGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 476
 Db 588 GGTGAGGAGAGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 647
 Oy 477 TCTTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 536
 Db 648 TCGATTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 707
 Oy 537 TACTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596
 Db 708 TACTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 767
 Oy 597 GAAACAGAGAGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656
 Db 768 GAAACAGAGAGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 827
 Oy 657 GGTGAGGAGAGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 716
 Db 828 GGTGAGGAGAGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 887
 Oy 717 AAACATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 775
 Db 888 AAACATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 947
 Oy 776 TTAATAGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 835
 Db 948 TGAATTTAGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1005
 Oy 836 AATTGGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 895
 Db 1006 AATTGGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1042
 Oy 896 ACATTTGATTTCAAAATGTTTCTTCTGATGCTGTTTATTTTCAAAATGTTTCTTCT 955
 Db 1043 AAATTTGATTTCAAAATGTTTCTTCTGATGCTGTTTATTTTCAAAATGTTTCTTCT 1102
 Oy 956 ATCAAAATGTTTCAAAATGTTTCTTCTGATGCTGTTTATTTTCAAAATGTTTCTTCT 1015
 Db 1103 GGCATGATTTCAAAATGTTTCTTCTGATGCTGTTTATTTTCAAAATGTTTCTTCT 1160
 Oy 1016 CTTTGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1075
 Db 1161 CTTTGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1216
 Oy 1076 TATGAGTATGTTCAAAATGTTTCTTCTGATGCTGTTTATTTTCAAAATGTTTCTTCT 1135
 Db 1217 TAGATTTATGTTCAAAATGTTTCTTCTGATGCTGTTTATTTTCAAAATGTTTCTTCT 1276

Oy 1136 GTACAAAAAGAAATATATA 1155
 Db 1277 GCTAAAGAGAGAGACTGTA 1296

RESULT 9
 HSCRISP21
 LOCUS HSCRISP21 1406 bp mRNA linear PRI 12-APR-1996
 DEFINITION H.sapiens mRNA for cysteine-rich secretory protein-2/type I.
 ACCESSION X95239
 VERSION X95239.1 GI:1262816
 KEYWORDS CRISP-2 gene; cysteine-rich secretory protein-2/type I.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 1406)
 Kratzschmar, J., Haendler, B., Eberspaecher, U., Roosterman, D., Donner, P. and Schlenning, W.D.
 The human cysteine-rich secretory protein (CRISP) family. Primary structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3
 Eur. J. Biochem. 236 (3), 827-836 (1996)

JOURNAL
 MEDLINE
 PUBMED
 96270732
 8665901
 2 (bases 1 to 1406)
 Haendler, B.
 Direct Submission
 Submitted (18-JAN-1996) B. Haendler, Scherling AG, ICMB, S109/517, 13342 Berlin, FRC
 Related sequence J04741.

COMMENT
 FEATURES
 source
 1..1406
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="testis cDNA library"
 241..972
 /gene="CRISP-2"
 241..972
 /gene="CRISP-2"
 /codon_start=1
 /product="cysteine-rich secretory protein-2/type I"
 /protein_id="CA64526.1"
 /db_xref="GI:1262817"
 /db_xref="SWISS-PROT:P16562"
 /translation="MALPVLFLVTLPLSPAPGKDPATFALLTLOLOVREIVNKH
 NELKRAVSPASNNLKMESREVTNORMAKCTLOHSDERKTSRGENLYMS
 DP7SMSSAIOHWDEILDYGVGPKSNAYGVHTOLVYSTVOVCGIAYCENODS
 LKTVICQYCPAGNNMKNRPVQGGFPCACPDGDKGLCTNSCYODLNSCDLSL
 NTPACERHELLEKRCATCLCNKLY"
 241..300
 /gene="CRISP-2"
 polyA_signal 1362..1367
 BASE COUNT 458 a 293 c 289 g 366 t
 ORIGIN

Query Match 24.0%; Score 513.6; DB 9; Length 1406;
 Best Local Similarity 68.7%; Pred. No. 8.5e-85;
 Matches 797; Conservative 0; Mismatches 324; Indels 39; Gaps 5;

Oy 1 TGATAAACAAATATTCTTATCTGCTGGAACCA---CTGCATGATTTATCCCA 56
 Db 193 TAATAAGATAGATATTCTATCTGCTGGAACCAACCAATTTCCGATGGCTTACTA 252
 Oy 57 GTCGCTGTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 116
 Db 253 CCGGCTGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309
 Oy 117 CCGGCTTATTGCTTGTGTTAACCAACCAAGAGAGAGAGATGTAATAG 176
 Db 310 CCGGCTTATTGCTTGTGTTAACCAACCAAGAGAGAGAGATGTAATAG 369
 Oy 177 CACATGAGATGAGAGAGAGATATCTCCCTGCGAGAAACATCTGAGATGAGATG 236

```

Db 370 CACATGATAGTAAAGAACGACTCTCTCCACCTGCCAGTAAATGCTAAAGATGGAATGG 429
Oy 237 AACAAAGAGCGTCAGCAAAATGCCCAAAAGTGGGCAAAACCAAGTCAATTAACAGACAGCT 296
Db 430 AGCAGAGAGGTAAACAAGATGCCCAAAAGTGGGCAAAACCAAGTCAATTAACAGACAGCT 489
Oy 297 AACCCAAAGATGATGACAAAGTCTAAATGTGTGAGAAATCTCTACATGTGCAGTGGC 356
Db 490 GATCCAGAGACCGCAAAACCAAGTCAATGATGTGGAATCTCTATGTCTCAAGTGC 549
Oy 357 CCCAGCTCATGTGTCAACCAAGTCAATGATGTGTGATGATGATGATGATGATGATGATGAT 416
Db 550 CCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 609
Oy 417 GGTGAGGGCCAAAGACTCCCAAGCAGTGTGTGACATTAATACAGAGTGTGTGTGAC 476
Db 610 GGTGAGGACCAAAAGAGTCCCAATGCTGTGTGACATTAATACAGAGTGTGTGTGAC 669
Oy 477 TCTCATACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 536
Db 670 TCGACTTACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 729
Oy 537 TACTATGTTGCCAATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 596
Db 730 TACTATGTTGCCAATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 789
Oy 597 GAACAAGAGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 656
Db 790 CAACAAGAGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 849
Oy 657 GGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 716
Db 850 ACTTGTGCTATCAAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 909
Oy 717 AACATCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 775
Db 910 GAACATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 969
Oy 776 TTAATACCATTTACACACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 835
Db 970 TGAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1027
Oy 836 ATTTGGCATCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 895
Db 1028 -----GCCACATPACCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1064
Oy 896 ACATTGTGATTAATGTTTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 955
Db 1065 AATATTGATTTAAAGAGCAATGCAATCTTCTCCCCAGATCAACAGAAATCACTTTGA 1124
Oy 956 ATACAAATGTTAAAAAGCAAAATCTATACAAACAACTTTGATTTTATATATAA 1015
Db 1125 GGCATGATTTTCAAAAGAGCATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1182
Oy 1016 CTTTGTGATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1075
Db 1183 -----TTTACTGCTTTATACAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1238
Oy 1076 TATGACTAAGTTCATTAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1135
Db 1239 TAGGATTTGTCCTACTAGAGCTTTGGATCAAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1298
Oy 1136 GTACAAAAAGAACATATA 1155
Db 1299 GCTAAGAAGAGACTGTA 1318

```

```

RESULT 10
ECRISP3 ECRISP3 1295 bp mRNA linear MAM 03-FEB-1998
LOCUS DEFINITION Equus caballus mRNA for cysteine-rich secretory protein-3.
ACCESSION AF001400

```

```

VERSION AJ001400.1 GI:2388784
KEYWORDS CRISP-3; cysteine-rich secretory protein-3.
SOURCE horse.
ORGANISM Equus caballus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
1 (bases 1 to 1295)
AUTHORS Schambony, A.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-1997) Schambony A., Institut fuer
Reproduktionsmedizin, Tierärztliche Hochschule Hannover, Bunteweg
15, D-30559 Hannover, GERMANY
REFERENCE 2 (bases 1 to 1295)
AUTHORS Magdalenio, L., Gassel, M., Vree, J., Schambony, A.M., Urbanke, C.,
Raida, M., Topfer-Petersen, E. and Calvete, J.J.
TITLE Biochemical and conformational characterization of HSP-3, a
stallion seminal plasma protein of the cysteine-rich secretory
protein (CRISP) family
JOURNAL FEBS Lett. 420 (2-3), 179-185 (1997)
MEDLINE 98119394
PUBMED 9459306
FEATURES
source Location/Qualifiers
1..1295
/organism="Equus caballus"
/db_xref="taxon:9796"
/sex="Male"
/rissue_type="ampulla"
/dev_stage="adult"
59..796
/gene="crisp-3"
59..796
/gene="crisp-3"
/codon_start=1
/product="cysteine-rich secretory protein-3"
/protein_id="CAA04729.1"
/db_xref="GI:2388785"
/db_xref="E:2388785"
/translacion="MALP.PVL.FLA.VL.I.P.P.P.A.S.P.O.P.G.P.A.L.S.I.T.S.E.V.O.K.E.I.V.N.K
H.N.D.L.R.T.V.S.P.L.A.S.N.M.L.K.W.O.M.B.S.K.T.I.N.M.O.N.A.N.K.L.L.H.S.K.E.D.A.V.T.M.K.G.E.L.P.H
S.S.I.P.N.S.D.A.I.O.N.H.D.E.V.H.F.R.Y.G.V.P.T.P.N.A.V.G.H.Y.O.V.V.S.S.Y.R.V.G.C.I.A.T.C.P.K.O
G.T.L.K.Y.V.V.C.O.R.C.P.A.G.N.V.Y.N.K.I.N.T.P.E.O.G.T.P.C.A.R.G.N.C.D.N.G.I.C.T.S.C.E.B.L.V.S.N.C.D.S
L.K.I.N.G.C.H.E.L.L.E.N.C.K.T.T.C.O.C.E.N.R.I.Y"
59..124
sig_peptide
polyA_signal 1262..1267
BASE COUNT 428 a 257 c 245 g 365 t
ORIGIN
Query Match 23.6%; Score 505.8; DB 4; Length 1295;
Best Local Similarity 70.5%; Pred. No. 2.4e-83;
Matches 803; Conservative 0; Mismatches 297; Indels 39; Gaps 8;
Oy 4 TGAACAAATACTTCACTCTGCTGTGGAACCACTGCAATGACATTAATCCAGTCTGT 63
Db 24 TAAAGAGATGCTTCACTGCTGTGGAACCA---ACATGGCATTAATACAGTGTGCG 80
Oy 64 TGTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 123
Db 81 TGTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 137
Oy 124 TTAAGTCTTTGTTAAACCAACCAACAGTGAAGGAGATGTGATTAACCAATG 183
Db 138 TTGCTCTTTTGTCAATCACCACCAAGCAAGTCCAAAAGAGATTTGTAATTAACCAATG 197
Oy 184 AACTGAGAGAGAGATATCTCCCTGTCGAGAAACATGCTGAAGATGAGATGAGAACAAAG 243
Db 198 ACCTAAGAGAGACAGTCTCTCCACTTGCCAGTAAATGCTAAAGATGAGATGAGAACAAAG 257
Oy 244 AGGCTGACGAATGCCCCAAAGTGGCAAAACGATGCAATTAACAGACAGTAAACCAA 303
Db 258 AGACAGCAACCAATGCCCCAAAGTGGCCAAAGAGGCTTCTCCAAACAGTAAAGCAG 317
Oy 304 AGAGTCG---AATGACAAAGTCTAAATGTGTGAGAAATGCTCAATGTCAGTCAAGTGC 360

```


QY	936	ATTTCACAAATATTTTCATACCAAAAGTGTAAAAAGAAACAAATATATACCAACAC	995
Db	1076	TTGATACATTAATATCTTTTCAGACACGCCATTTAAAAAGTTGTCATTTATATACCAAC	1135
QY	996	TTTGATTTTATATATAAACTTTGATTTAAATTTACTGATTAATTTAGGTGAGAA	1055
Db	1136	TTGGCAATTTGA---CCAACTTGGTGGCTTTAAATATAGTGAACGAAATCAAGTGAAGA	1192
QY	1056	TTTGAAGTTTATATTCATATATGACTAAGTTCACTAAACCCTGATTTGAAAGTGAA	1115
Db	1193	TTTGAAGTTGCACCAATATATGCGTCATTTACCTCTGCAACTTTGAAAAAAGAGAAAC	1252
QY	1116	TTATGTCTTCAGACAAATATGACAA	1141
Db	1253	TATTTGTGTCCTAAATTAACCTGCAA	1278
RESULT 13			
LOCUS	AB009662	1280 bp	mRNA
DEFINITION	Rattus norvegicus mRNA for testis specific protein, complete cds.		linear
ACCESSION	AB009662		
VERSION	AB009662.1	GI:3374579	
KEYWORDS	testis specific protein.		
SOURCE	Rattus norvegicus (strain:Donryu) 20 days after birth male Testis		
ORGANISM	Spermatogenic cells cDNA to mRNA, clone:Tpx-1.		
	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus		
REFERENCE	1 (bases 1 to 1280)		
AUTHORS	Meda,T., Sakashita,M., Ohba,Y. and Nakanishi,Y.		
TITLE	Molecular cloning of the rat Tpx-1 responsible for the interaction		
	between spermatogenic and Sertoli cells		
JOURNAL	Biochem. Biophys. Res. Commun. 248 (1), 140-146 (1998)		
REFERENCE	98340864		
	2 (bases 1 to 1280)		
AUTHORS	Nakanishi,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-DEC-1997) Yoshinobu Nakanishi, Kanazawa University,		
	Faculty of Pharmaceutical Sciences, 13-1 Takara-machi, Kanazawa,		
	Ishikawa 920, Japan (Tel:076-234-4460, Fax:076-234-4480)		
FEATURES	Location/Qualifiers		
source	1..1280		
	/organism="Rattus norvegicus"		
	/strain="Donryu"		
	/db_xref="taxon:10116"		
	/clone="Tpx-1"		
	/sex="male"		
	/cell_type="Spermatogenic cells"		
	/tissue_type="Testis"		
	/dev_stage="20 days after birth"		
	78..809		
	/codon_start=1		
	/product="testis specific protein"		
	/protein_id="BAA32029.1"		
	/db_xref="GI:3374580"		
	/translation="MAMFQVLAFLPFAVLPLPTEGKDDPFAITLTTNOIOVEPIIAK		
	HNELRQVSPGSSNLIKMEKNVQGAANAKQKANNICIEHSTEDRKINICGELNHS		
	TDPRSRTVIOSTRLENENFVVGGAAPNSAVGHYTOLVWSSFKVCGVAYCPNDOT		
	LKFIYVCHPCMGNNVAKKSTPYHGQPCASCPNMCNDGICLTSNCFDELLSNESIK		
	SSAGCKHEHLAKKEAECALCEDEKHK"		
	1280		
	/note="20 a nucleotides"		
BASE COUNT	416 a	263 c	259 g
ORIGIN			342 t
Query Match	17.2%;	Score 367.8;	DB 10;
Best Local Similarity	61.8%;	Pred. No. 5.9e-58;	Length 1280;
Matches 687;	Conservative 0;	Mismatches 387;	Indels 37;
		Gaps 5;	
37	CTGCATGACATTTATCCCACTGCTGTTGTTCTGGTTCGTGGGCTGCTCCATCTTTTC	96	

Db	73	CAGCATTGGCTGGTTCAGGATGCTGTTTGTCTTGTGCTGTGCTGATACA	129
Qy	97	CAGCAATGGAAGATTAAGATCCGCCCTTTACTGCTTTGTTAACCCACCAACAAGTGC	156
Db	130	CACCCACAGAAAGAAAGGATCCAACTCTGCTCTACTTTGACMACCAACCAATACAAATTTC	189
Qy	157	AAAGGAGATTGTGAATTAAGCACAAATGAATGAGAGAGACAGTATCTCCCTCCGACGA	216
Db	190	AAAGGAGATATATAGCTTAACACAAATGAATGAGAGACAAAGTTACCCCTCCGACGA	249
Qy	217	ACATCTGAGACATGAATGGAACAAAGGCTGCAGCAATGCCCAAAAGTGGGCAACC	276
Db	250	ACATCTCTAAATGGAATGGAACCTGAACACACAGCAAAATGCTCAAAAGTGGCTAATA	309
Qy	277	AGTGAATTACAGACACAGTAACCCAAAGGATGGAAATGACAAATGCTAAATGGTGAGA	336
Db	310	ACTGATTTTGAACACAGTAGTACAAACCGAAATCAATATTAATTAATGGCGAGA	369
Qy	337	ATCTCTACATCTCAAGTGGCCCCAGCTCATGCTCACAGCAATCCAAAGCTGGTTGATG	396
Db	370	ATCTCTAATATGTCAAGTACAGACCTCATCATCTGTGAGAAACCGTAATTAACAACTGATGAG	429
Qy	397	AGTACATGATTTTGACTTTGGTGTAGGCGCAAAAGACTCCCAACGCAAGTGGGCAATT	456
Db	430	AAATGAAACCTGTGTTTGGCGCTAGAGACTA...-ACCCAATTCGCGTGTGCGCACT	486
Qy	457	ATACACAGTGTGGTGTACTCTTCATACCTGCTGGATGTGGAATGGCTACTGCCA	516
Db	487	ACACTCAGCTTGTTGGTATTCATCTTTCAAATTTGGATGTGAGTGTGCTTACTGCCA	546
Qy	517	ATCAAAAGTTCTAAATTAATACTATGTTTCCAAATATGTGCTGCTGTAATGGGCTA	576
Db	547	ATCAGAGTACCCCTAAATACTCTTATGTTGGCCATTACTGCTCTATGGGTAAACAACGTGA	606
Qy	577	ATAGACTATATGTCCTTATGAACACAGGACACTGTGCGAGTTGGCCCAAGTAACATG	636
Db	607	TGAAAAGAGTACCCCATATCATCMAAGGACACTTGTGTAGTGTGTCCCAATTAATG	666
Qy	637	ACGATGCACTATGCAACCAATGGTTGCAAGTAGCAAGATCTATAGTAACCTGTAAAGTT	696
Db	667	ATAATGATTTGTGCAACCAATAGCTGTGATTTTGAATTTACTTATGTAACCTGTGAATCT	726
Qy	697	TGAAGCTCACTTAACCTGTAAACATCAGTGTGTGACGGACAGTTGCAAGCATCTCTCA	756
Db	727	TGAAGACTTACGCGCTGTAAACATGAGTGTCTCAAGGCCAAAGTGTGAGGCTACTTCC	786
Qy	757	ATTGTTCAAAACAGATTATTAATATAGCATATGCAACAGGATAGAGGCTATGACAGAG	816
Db	787	TATGTGAAGACAAATTCATTAAACATGAGCCAGGCTGCACATGACAGACTATATGAGAG	846
Qy	817	AGTACAGTTATCTACTAGTAAATTTGGCATCTACTTACATTTAACATATACAGTACAGAA	876
Db	847	GGGTACAGACTAGTTGAGCATGAGACA...-GGGAAAA	880
Qy	877	TTGTAGCATGTTTGAATACACATTTATTTCAATGTTTCTTCTGTGATCTGCTTTTA	936
Db	881	CCATATGAGAGATGATGAAAACAGTGCATCCCAATATACAAAGGCTTCTTCTCTCTGAT	940
Qy	937	TTTACAAAATATTTTTCATACAAATGGTTAAAGAAAGAAACAAATCTATAACAACACT	996
Db	941	TTATATGAAGATGCTTCTCATACAGCATTAAGAAAG--GTCTATTTAGGATTAACAAC	998
Qy	997	TTGATTTTATATATTAACCTTTGTGATTTAAATTTACTGAATTTAATTTAGGGTGAAT	1056
Db	999	CTGATTTTGA...-CCAACTTCTCTCTTCAATATGATGAGGCAACCAAGTGGAGAT	1055
Qy	1057	TTTGAAGTTGTATTCATATGACATAGTACATTAAGTCAATAAACCTGTGATGAAGTGAAT	1116
Db	1056	TTTGAAGTTGTATTCATATGACATAGTACATTAAGTCAATTAAGTGAAGTGAAT	1115
Qy	1117	TATGTTCTGAGAACAAATGTACAAAAGAA	1147

AUTHORS Charast, N.J., Joseph, D.R., Wilson, E.M. and French, F.S.
 TITLE Molecular cloning of complementary deoxyribonucleic acid for an androgen-regulated epididymal protein: sequence homology with metalloproteins
 JOURNAL Mol. Endocrinol. 2 (10), 999-1004 (1988)
 MEDLINE 89039913
 PUBMED 2460753

FEATURES
 source location/Qualifiers
 1..1498
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue-type="epididymus"
 /dev_stage="adult"
 1..1498
 /gene="AEG"
 <1..1498
 /gene="AEG"
 /product="alternative mRNA"
 join(<1..918,1457..1493)
 /gene="AEG"
 /product="alternative mRNA"
 11..851
 /gene="AEG"
 /codon_start=1
 /product="epididymal glycoprotein"
 /protein_id="AAB59716.1"
 /db_xref="GI:202773"
 /translation="MALMLYLFLAALVPSLDDTTDEMNDLENLSTKLSVOEEL
 INKNQRLRTVSPSGSDLRVEMDHDAVYNAOKVANCITVHSPLOHRTTLKGENL
 FMANYPASMSVYIDWYDESLDFVFGPKVGVKVAHTQVYVNSPIVACGAEEDP
 DQPKYFVCHVGCGNVRILVSPYRGEPCDPCPNCEGDLCTNCEIEDNTSNCS
 DLKMSCDPLKEGCRASCFCEDKIH"
 1476..1477
 /gene="AEG"
 /note="at in long clone; aat in short clone"
 /replace="aat"
 1490
 /gene="AEG"
 /note="t in long clone; c in short clone"
 /replace="c"
 variation
 BASE COUNT 464 a 289 c 271 g 474 t
 ORIGIN

Query Match 16.98; Score 362.2; DB 10; Length 1498;
 Best Local Similarity 62.38; Pred. No. 6,1e-57;
 Matches 713; Conservative 0; Mismatches 403; Indels 29; Gaps 8;

```

OY 11 AATCTTCATCCTGCTCGAAGAACACATGACATGATTATCCACAGCTGTTGTTCT 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 AATATCTCATCTCTCT-GAAATATGAAACATGGCATTAATGTTAGCTGTTGTTCT 139
OY 71 GGTTCGTGGCTGCTTCATCTTTCCAGCAAA-----TGAAGATAAGATCCGCTTT 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 140 GGTGCTGTATGCGACCATCTCTCTTCAGATATACACTGATGATGGATAGAGACT 199
OY 125 TACTGCTTTGTTAAACCAACCAACAGTGCAGAGAGATTGTGAATTAACACATGA 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 TGAGAAATTTGCAACCACTAACTGTCAGTCAAGAAAGATCATTAACCAACCA 259
OY 185 ACTGAGAGACAGATATCTCCCTGCCAGAAACATGCTGAAGATGGAATGGAAACA 244
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 ATTGAGACGAAGGTTCTCGTCTGTAGACTTACTAAGAGTGAATGGACATGA 319
OY 245 GGTGCAAGCAATGCCCAAAAGTGGCAAAACAGTCAATTACAGACAGTAAACCA 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 TGTATGTGAACGCTGGAATAGGCAACAGGTGCAATTACATCACTACATCTCTACA 379
OY 305 GGATCGAATGACAAAGTGAATGTGTGAGAAATCTTACATGTCAAGTGCCTCCAGCTC 364
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 ACACAGGACCAACATTAATAATGTGTGAGAAATTTGTTACATGCAATACCTGCATC 439
OY 365 ATGTCACAAACCAATCCAAAGCTGTTGATGATACATGATTTTGTGCTTGGTGAG 424

```

```

Db 440 GTGGTCTCTGTGAATCCAGATGGTATGATGAATCCCTTGATTTTGTGTTGG 499
OY 425 GCCAAGACTCCCAACGAGAGTGTGACATTAATACAGAGTGTGTGCTCTTCATA 484
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 500 CCCAAAAAAGTGTGTTAAAGTGGACACTATATCTAGGTGTTTGAATTCACTTT 559
OY 485 CCTGCTGATGTGGAATGGCTTACTGTCCCAATCAAAAAAGTCTTAATACTACTAT 544
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 560 CCGTGTGATGTGAGAGTGTGATGATGCCCTGACCAACCA--TGAATATCTTTATGT 616
OY 545 TTGCCAATTTGCTCTGCTGTGAATTTGGGCTATATGATATATGCTTATGAACA 604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 617 TTGTACTATTTGCTGTGTCGCAATTAATGTAAGAACTATACCTCACTACAGAA 676
OY 605 AGCACTTGTGCAAGTTCGCCAGATTAAGTGTGACATGAGTATGACCAATGATGCA 664
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 677 AGAAGCTTGTGACAGTTCCTGCTGTAATTTGTAAGATGGCTGTGACCAATAGTTG 736
OY 665 GTACGAGATCTTATTAATTAATCTGTAAGAGTTGAAGTCAATTAACCTTAACAT 724
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737 ATATGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 796
OY 725 GTTGTGAGGAGACAGTTCAGAGGATCTCTGCAATTTGTTCAACAGCATTTTAAT 784
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 797 ACTTCTTAAGAAAGTTCAGAGGATCTCTGTAATTAATTAATTAATTAATTAAT 856
OY 785 CATACACACGAGTGGCTATGTAAGAGAGAGTCAATTAATTAATTAATTAATTAAT 844
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 857 CAGTCCNA-TATACAGACCATGATAGAAAAAGAAAAAATCCCTGACTTACTTAT 915
OY 845 CTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 904
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 916 GTCC-----CACCAAAATATGAGTTGTAGTACTGAATTAATTCAAATGT 965
OY 905 TTCAAGTCTTTCTGCGATCTGCTTTATTAATTAATTAATTAATTAATTAATTA 964
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 966 AAGATTTCTGTTCTCTCTCTATTTCTCTATTTTGCATTAATTAATTAATTAAT 1025
OY 965 GTTAAAGAAACAAATCTATTAACACACTTGGATTTTATTAATTAATTAATTAAT 1024
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1026 TTTTAAATTAACAAA--TCAATACACCTTTTGAAGTCCCATATGAATCTGTGA - 1080
OY 1025 TTAATTTACTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1084
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1081 -CACATTTATGAAATCAATATATCCACAGATTAAT--ATATTATTTGTCTATGACTTA 1137
OY 1085 GTTCACTAAACCTCGATTTGAAGTGAATTAATTAATTAATTAATTAATTAATTA 1144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1138 AGTCACTAAATCTCTGCTGTAATTAATTAATTAATTAATTAATTAATTAATTA 1197
OY 1145 GAACA 1149
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1198 GAACA 1202

```

Search completed: March 14, 2003, 02:07:40
 Job time : 6413 secs